# Non-quantum Entanglement and a Complete Characterization of pre-Mueller and Mueller Matrices in Polarization Optics

B. Neethi Simon

*Department of Applied Mechanics, IIT Madras, Chennai 600 036*

Sudhavathani Simon

*Department of Computer Science, Women's Christian College, Chennai 600 006*

N. Mukunda

*Centre for High Energy Physics, Indian Institute of Science, Bangalore 560 012*

F. Gori and Massimo Santarsiero

*Dipartimento di Fisica, Universitá Roma Tre and CNISM, Vila della Vasca Navale 84, I-00146 Rome*

Riccardo Borghi

*Dipartimento di Electronica Applicata, Universitá Roma Tre and CNISM, Vila della Vasca Navale 84, I-00146 Rome*

R. Simon

*The Institute of Mathematical Science, Tharamani, Chennai 600 113*

The Mueller-Stokes formalism which governs conventional polarization optics is formulated for plane waves, and thus the only qualification one could demand of a $4 \times 4$ real matrix $M$ in order that it qualifies to be the Mueller matrix of some physical system is that $M$ should map $\Omega^{(\mathrm{pol})}$, the positive cone of Stokes vectors, into itself. In view of growing current interest in the characterization of partially coherent partially polarized electromagnetic beams, there is need to extend this formalism to such beams wherein the polarization and spatial dependence are generically inseparably intertwined. This inseparability or non-quantum entanglement brings in additional constraints that a pre-Mueller matrix $M$ mapping $\Omega^{(\mathrm{pol})}$ into itself needs to meet in order that it is an acceptable physical Mueller matrix. These additional constraints are motivated and fully characterized.

*OCIS codes*: 260.5430, 270.5585, 260.2110, 120.5410, 230.5440

## I. INTRODUCTION

Entanglement is traditionally studied almost exclusively in the context of quantum systems. However, this notion is basically kinematic in nature, and so is bound to present itself whenever and wherever the state space of interest is the tensor product of two (or more) vector spaces. The vectors of the individual spaces, and hence (tensor) products of such vectors, will be expected to possess identifiable physical meaning. Polarization optics of paraxial electromagnetic beams happens to have precisely this kind of a setting, and so one should expect entanglement to play a nontrivial role in this situation. It turns out that entanglement in this non-quantum setup is not just a matter of academic curiosity: we shall show in this paper that *this non-quantum entanglement helps in resolving a fundamental issue in classical polarization optics*. It will appear that this issue could not have been resolved without explicit consideration of entanglement.

A paraxial beam propagating along the positive $z$-axis is completely determined in terms of the transverse components of the electric field specified throughout a transverse plane $z =$ constant as functions of the transverse variables $(x, y) = \boldsymbol{\rho}$. If these components are independent of the transverse coordinates, then the situation corresponds to a plane wave propagating along the $z$-axis. The traditional Mueller-Stokes formalism in terms of Stokes vector $S$ and Mueller matrix $M$, describing respectively the beam and the optical system, presumes essentially this kind of situation wherein the spatial degree of freedom can be safely left out of consideration, the focus being on the polarization degree of freedom [1, 2, 3].

Recent years have witnessed an enormous interest in partially polarized partially coherent electromagnetic beams [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18] and hence there is a need to extend the Mueller-Stokes formalism to such beams. Given a $4 \times 4$ real matrix $M$, it should map $\Omega^{(\mathrm{pol})}$, the positive cone of Stokes vectors, into itself in order

that it could be the Mueller matrix of some physical system. *Within the conventional formalism, this seems to be the only qualification that can be demanded of $M$.* In a partially coherent partially polarized beam, polarization and spatial dependence happen to be inseparably intertwined. This inseparability or non-quantum entanglement brings in additional constraints that a pre-Mueller matrix $M$ mapping $\Omega^{(\text{pol})}$ into itself needs to meet in order that it is a physically acceptable Mueller matrix.

The principal purpose of the present work is to motivate these constraints and characterize them fully. The next two Sections of the paper act as preparation towards this end. We begin in Section 2 by recounting the Mueller-Stokes formalism as it applies to plane waves. This is then extended in Section 3 to paraxial electromagnetic beams, and the role of (non-quantum) entanglement between polarization and spatial modulation is rendered transparent. These two Sections equip us with all the tools needed to formulate in Section 4 the additional physical constraints on a pre-Mueller matrix $M$, arising as consequence of entanglement. Our final result is formulated in the form of a necessary and sufficient condition, and a simple illustrative example is treated in some detail for illustration of the nature of these further constraints. And we conclude in Section V with some further remarks.

## II. POLARIZATION OPTICS OF PLANE WAVES

For a plane wave whose propagation direction is along the (positive) $z$-axis perpendicular to the $(x, y)$-plane, the $x$ and $y$ components $E_1$, $E_2$ of the electric field are independent of the transverse-plane coordinates $\boldsymbol{\rho}$, and can be arranged into a (numerical) *column vector*

$$\boldsymbol{E} \equiv \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \in \mathcal{C}^2. \tag{2.1}$$

We have suppressed, for convenience, a space-time dependent scalar factor of the form $e^{i(kz-\omega t)}$. While $\boldsymbol{E}^\dagger \boldsymbol{E} = |E_1|^2 + |E_2|^2$ is (a measure of) the intensity, the ratio $\gamma = E_1/E_2$ of the (complex) components, which ratio can be viewed as a point on the Riemann or Poincaré sphere $S^2$, specifies the polarization state. In particular, the signature of the imaginary part of $\gamma$ describes the handedness of the (generally elliptic) polarization.

In presence of fluctuations $\boldsymbol{E}$ acquires some randomness, and in this case the state of polarization is effectively described by the $2 \times 2$ *coherency or polarization matrix*

$$\Phi \equiv \langle \boldsymbol{E}\boldsymbol{E}^\dagger \rangle = \begin{bmatrix} \langle E_1 E_1^* \rangle & \langle E_1 E_2^* \rangle \\ \langle E_2 E_1^* \rangle & \langle E_2 E_2^* \rangle \end{bmatrix}, \tag{2.2}$$

where $\langle \cdots \rangle$ denotes ensemble average. The coherency matrix is hermitian, $\Phi^\dagger = \Phi$, and positive semidefinite, $\boldsymbol{V}^\dagger \Phi \boldsymbol{V} = \text{tr}\,(\Phi \boldsymbol{V}\boldsymbol{V}^\dagger) \geq 0$, $\forall \boldsymbol{V} \in \mathcal{C}^2$. This positivity property may be denoted simply $\Phi \geq 0$. Hermiticity and positivity are the *defining properties* of $\Phi$: every $2 \times 2$ matrix obeying these two conditions is a valid coherency matrix, and represents some polarization state. Since $\Phi$ is a $2 \times 2$ matrix, the positivity condition takes the simple scalar form

$$\begin{aligned} \text{tr}\,\Phi &> 0, \\ \det\,\Phi &\geq 0. \end{aligned} \tag{2.3}$$

It is clear that the intensity corresponds to $\text{tr}\,\Phi$. Fully polarized light (pure states) corresponds to $\det\,\Phi = 0$, and partially polarized or mixed states correspond to $\det\,\Phi > 0$.

*Typical systems of interest in polarization optics are transversely homogeneous*, in the sense that their action is independent of the coordinates $(x, y)$ spanning the transverse plane in which the system lies. If such a system is deterministic and acts linearly at the field amplitude level, it is described by a complex $2 \times 2$ numerical matrix $J$ called the Jones matrix of the system:

$$\begin{aligned} J &= \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{12} \end{bmatrix} : \quad \boldsymbol{E} \to \boldsymbol{E}' = \begin{bmatrix} E_1' \\ E_2' \end{bmatrix} = J\boldsymbol{E} \\ &\Leftrightarrow \quad \Phi \equiv \langle \boldsymbol{E}\boldsymbol{E}^\dagger \rangle \to \Phi' = \langle \boldsymbol{E}'\boldsymbol{E}'^\dagger \rangle = J\Phi J^\dagger. \end{aligned} \tag{2.4}$$

It is clear that Jones systems map pure states ($\det \Phi = 0$) into pure states.

Since $\Phi$ is hermitian, it can be conveniently described as *real* linear combination of the four orthogonal hermitian matrices $\tau_0 = 1_{2\times 2}$, $\tau_1 = \sigma_3$, $\tau_2 = \sigma_1$, $\tau_3 = \sigma_2$:

$$\Phi = \frac{1}{2}\sum_{a=0}^{3} S_a \tau_a \iff S_a = \text{tr}(\tau_a \Phi); \ \text{tr} \, \tau_a \tau_b = 2\delta_{ab}. \tag{2.5}$$

The reason for choosing the $\tau$-matrices, a permuted version of the Pauli matrices rather than the Pauli $\sigma$-matrices themselves, is to be consistent with the optical convention that the circularly polarised states, the eigenstates of $\sigma_2$, be along the 'third' axis (polar axis) of the Poincaré sphere. The intensity equals $S_0 = \text{tr} \, \Phi$. The expansion coefficients $S_a$ are the components of the *Stokes vector* $S \in \mathbf{R}^4$. Note that $\tau_3^* = -\tau_3$ and $\tau_a^* = \tau_a$ if $a \neq 3$.

While hermiticity of $\Phi$ is equivalent to reality of $S \in \mathbf{R}^4$, the positivity conditions $\text{tr} \, \Phi > 0$, $\det \Phi \geq 0$ read, respectively,

$$\begin{aligned} S_0 &> 0, \\ S_0^2 - S_1^2 - S_2^2 - S_3^2 &\geq 0. \end{aligned} \tag{2.6}$$

Thus, *permissible polarization states correspond to the positive light cone and its interior* (solid cone). Pure states live on the surface of this cone. As suggested by this light cone structure, the *proper orthochronous Lorentz group* $SO(3,1)$ plays quite an important role in polarization optics [19, 20].

Under the action of a deterministic or Jones system $J$ described in (2.4) the elements of the output coherency matrix $\Phi'$ are obviously linear in those of $\Phi$. This, in view of the linear relation (2.5) between $\Phi$ and $S$, implies that under passage through such a system the output Stokes vector $S'$ and the input $S$ will be linearly related by a $4 \times 4$ real matrix $M(J)$ determined by $J$:

$$J: \ S \to S' = M(J)S. \tag{2.7}$$

We may call $M(J)$ the Mueller matrix of the Jones system $J$. It is known also as a Mueller-Jones matrix to emphasise the fact that it is constructed out of a Jones matrix. While $\Phi = \langle \boldsymbol{E}\boldsymbol{E}^\dagger \rangle$ is a $2 \times 2$ matrix, the tensor product $\tilde{\Phi} \equiv \langle \boldsymbol{E} \otimes \boldsymbol{E}^* \rangle$ is a four-dimensional column vector associated with $\Phi$:

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\Phi}_0 \\ \tilde{\Phi}_1 \\ \tilde{\Phi}_2 \\ \tilde{\Phi}_3 \end{bmatrix} \equiv \begin{bmatrix} \langle E_1 E_1^* \rangle \\ \langle E_1 E_2^* \rangle \\ \langle E_2 E_1^* \rangle \\ \langle E_2 E_2^* \rangle \end{bmatrix} = \begin{bmatrix} \Phi_{11} \\ \Phi_{12} \\ \Phi_{21} \\ \Phi_{22} \end{bmatrix}. \tag{2.8}$$

This idea of going from a pair of indices, each running over 1 and 2, to a single index running over 0 to 3 and vice versa can often be used to advantage to associate with any $2 \times 2$ matrix $K$ a corresponding column vector $\tilde{K}$ with $\tilde{K}_0 = K_{11}$, $\tilde{K}_1 = K_{12}$, $\tilde{K}_2 = K_{21}$, and $\tilde{K}_3 = K_{22}$. The tensor product $J \otimes J^*$ is a $4 \times 4$ matrix:

$$\begin{aligned} J \otimes J^* &\equiv \begin{bmatrix} J_{11}J^* & J_{12}J^* \\ J_{21}J^* & J_{22}J^* \end{bmatrix} \\ &= \begin{bmatrix} J_{11}J_{11}^* & J_{11}J_{12}^* & J_{12}J_{11}^* & J_{12}J_{12}^* \\ J_{11}J_{21}^* & J_{11}J_{22}^* & J_{12}J_{21}^* & J_{12}J_{22}^* \\ J_{21}J_{11}^* & J_{21}J_{12}^* & J_{22}J_{11}^* & J_{22}J_{12}^* \\ J_{21}J_{21}^* & J_{21}J_{22}^* & J_{22}J_{21}^* & J_{22}J_{22}^* \end{bmatrix}. \end{aligned} \tag{2.9}$$

The transformation $\Phi \to \Phi' = J\Phi J^\dagger$ is thus equivalent to $\tilde{\Phi} \to \tilde{\Phi}' = J \otimes J^* \tilde{\Phi}$. Since $\tilde{\Phi}$ is related to the Stokes vector through

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Phi}_0 \\ \tilde{\Phi}_1 \\ \tilde{\Phi}_2 \\ \tilde{\Phi}_3 \end{bmatrix}, \tag{2.10}$$

it follows that

$$M(J) = A(J \otimes J^*)A^{-1}, \tag{2.11}$$

$A$ being the $4 \times 4$ matrix exhibited in Eq. (2.10); this matrix is *essentially* unitary: $A^{-1} = \frac{1}{2}A^\dagger$.

If $\det J$ is of unit magnitude, then $M(J)$ computed by this prescription is an element of $SO(3,1)$, the proper orthochronous group of Lorentz transformations; this was to be expected in view of the *two-to-one homomorphism* between $SL(2,C)$ and $SO(3,1)$. It follows that for any nonsingular $J$ the associated Mueller-Jones matrix $M(J)$ is $|\det J|$ times an element of $SO(3,1)$. *The prescription (2.11), though, applies to singular Jones matrices as well.*

### A.  Mueller matrices arising from Jones systems

A nondeterministic system is described *directly* by a Mueller matrix $M: S \to S' = MS$ and, by definition, such a Mueller matrix cannot equal $M(J)$ for any $2 \times 2$ (Jones) matrix $J$. Given a Mueller matrix $M$, how to test if it is a Mueller-Jones matrix for some $J$ or, equivalently, how to test if the system described by $M$ is a deterministic or Jones system? It turns out that this question which has received much attention [19, 21, 22, 23] has a simple and elegant solution [24].

We go over in some detail the construction underlying this solution, for it plays a central role in our analysis to follow. The sixteen $4 \times 4$ *hermitian* matrices $U_{ab} = \frac{1}{2}\tau_a \otimes \tau_b^*$, with $a$, $b$ independently running over the index set $\{0, 1, 2, 3\}$, form an orthonormal set or basis in the (vector) space of $4 \times 4$ matrices; these matrices are *unitary* and *self-inverses*:

$$\begin{aligned} U_{ab} &= \frac{1}{2}\tau_a \otimes \tau_b^* = U_{ab}^\dagger = U_{ab}^{-1}, \\ \mathrm{tr}\,(U_{ab}U_{cd}) &= \delta_{ac}\delta_{bd}, \quad a,b,c,d \in \{1,2,3,4\}. \end{aligned} \tag{2.12}$$

[Complex conjugation of the second factor of the tensor product $\tau_a \otimes \tau_b^*$ is suggested by the construct $J \otimes J^*$ in (2.11)]. Thus every (*hermitian*) $4 \times 4$ matrix can be written uniquely as a (*real*) linear combination of $\{\tau_a \otimes \tau_b^*\}$. Aa an important consequence of this fact we have:

**Proposition** 1: There exists a natural one-to-one correspondence between the set of all $4 \times 4$ real matrices and the set of all $4 \times 4$ hermitian matrices.

Indeed, a real matrix $M$ and the associated hermitian matrix $H_M$ are related in this simple manner:

$$H_M = \frac{1}{2}\sum_{a,b=0}^{3} M_{ab}\tau_a \otimes \tau_b^*. \tag{2.13}$$

Similarly, a hermitian matrix $H$ and the associated real matrix $M_H$ are related through

$$(M_H)_{ab} = \frac{1}{2}\mathrm{tr}\,(H\tau_a \otimes \tau_b^*), \quad a,b = 0,1,2,3. \tag{2.14}$$

It is clear that these relations are inverses of one another. We write these in more detail for later use:

$$H_M = \frac{1}{2}\begin{bmatrix} \begin{array}{c} M_{00}+M_{11} \\ +M_{01}+M_{10} \end{array} & \begin{array}{c} M_{02}+M_{12} \\ +i(M_{03}+M_{13}) \end{array} & \begin{array}{c} M_{20}+M_{21} \\ -i(M_{30}+M_{31}) \end{array} & \begin{array}{c} M_{22}+M_{33} \\ +i(M_{23}-M_{32}) \end{array} \\[1em] \begin{array}{c} M_{02}+M_{12} \\ -i(M_{03}+M_{13}) \end{array} & \begin{array}{c} M_{00}-M_{11} \\ -M_{01}+M_{10} \end{array} & \begin{array}{c} M_{22}-M_{33} \\ -i(M_{23}+M_{32}) \end{array} & \begin{array}{c} M_{20}-M_{21} \\ -i(M_{30}-M_{31}) \end{array} \\[1em] \begin{array}{c} M_{20}+M_{21} \\ +i(M_{30}+M_{31}) \end{array} & \begin{array}{c} M_{22}-M_{33} \\ +i(M_{23}+M_{32}) \end{array} & \begin{array}{c} M_{00}-M_{11} \\ +M_{01}-M_{10} \end{array} & \begin{array}{c} M_{02}-M_{12} \\ +i(M_{03}-M_{13}) \end{array} \\[1em] \begin{array}{c} M_{22}+M_{33} \\ -i(M_{23}-M_{32}) \end{array} & \begin{array}{c} M_{20}-M_{21} \\ +i(M_{30}-M_{31}) \end{array} & \begin{array}{c} M_{02}-M_{12} \\ -i(M_{03}-M_{13}) \end{array} & \begin{array}{c} M_{00}+M_{11} \\ -M_{01}-M_{10} \end{array} \end{bmatrix}. \tag{2.15}$$

This matrix in identical form was first presented in [24]. Of the sixteen $4 \times 4$ matrices $U_{ab}$, only $U_{20}$, $U_{21}$, $U_{30}$ and $U_{31}$ have *nonzero entries* at the $_{13}$ location, and this explains the entry $M_{20} - M_{21} - i(M_{30} - M_{31})$ for $(H_M)_{13}$. Written in detail, the relation (2.14) has the form

$$M_H = \frac{1}{2} \begin{bmatrix} H_{00} + H_{11} \\ + H_{22} + H_{33} & \begin{matrix} H_{00} - H_{11} \\ + H_{22} - H_{33} \end{matrix} & \begin{matrix} H_{01} + H_{10} \\ + H_{23} + H_{32} \end{matrix} & \begin{matrix} -i(H_{01} - H_{10}) \\ - i(H_{23} - H_{32}) \end{matrix} \\[1em] \begin{matrix} H_{00} + H_{11} \\ - H_{22} - H_{33} \end{matrix} & \begin{matrix} H_{00} - H_{11} \\ - H_{22} + H_{33} \end{matrix} & \begin{matrix} H_{01} + H_{10} \\ - H_{23} - H_{32} \end{matrix} & \begin{matrix} -i(H_{01} - H_{10}) \\ +i(H_{23} - H_{32}) \end{matrix} \\[1em] \begin{matrix} H_{02} + H_{20} \\ + H_{13} + H_{31} \end{matrix} & \begin{matrix} H_{02} + H_{20} \\ - H_{13} - H_{31} \end{matrix} & \begin{matrix} H_{03} + H_{30} \\ + H_{12} + H_{21} \end{matrix} & \begin{matrix} -i(H_{03} - H_{30}) \\ + i(H_{12} - H_{21}) \end{matrix} \\[1em] \begin{matrix} i(H_{02} - H_{20}) \\ + i(H_{13} - H_{31}) \end{matrix} & \begin{matrix} i(H_{02} - H_{20}) \\ - i(H_{13} - H_{31}) \end{matrix} & \begin{matrix} i(H_{03} - H_{30}) \\ + i(H_{12} - H_{21}) \end{matrix} & \begin{matrix} H_{03} + H_{30} \\ - H_{12} - H_{21} \end{matrix} \end{bmatrix}.$$

$$(2.16)$$

The entry $-i(H_{01} - H_{10}) + i(H_{23} - H_{32})$ for $(M_H)_{13}$ is explained by the fact that the nonzero entries of $U_{13}$ are at the $_{01}$, $_{10}$, $_{23}$ and $_{32}$ locations.

Fundamental to the structure of the Mueller-Stokes formalism of polarization optics is the following result:

**Proposition** 2: Given a $4 \times 4$ real matrix $M$, it is a Mueller-Jones matrix iff the associated hermitian matrix $H_M$ is a one-dimensional projection. That is iff $H_M = \tilde{J}\tilde{J}^\dagger$ for some (complex) four-dimensional column vector $\tilde{J}$. If $H_M = \tilde{J}\tilde{J}^\dagger$, then the $2 \times 2$ matrix $J$ associated with $\tilde{J}$ is the Jones matrix of the deterministic system represented by $M$.

In the original work [24] where this proposition was formulated and proved, this hermitian matrix $H_M$ was actually denoted $N$. However, in the present work we use instead the symbol $H$ to emphasise hermiticity, the most important property of this matrix.

Consider now a transformation which is a *convex sum* of Jones systems:

$$\Phi \to \Phi' = \sum_{k=1}^{n} p_k J^{(k)} \Phi J^{(k)\dagger}, \ p_k > 0, \ \sum_{k=1}^{n} p_k = 1. \tag{2.17}$$

This transformation may be realized by a set of $n$ deterministic or Jones systems $J^{(1)}$, $J^{(2)}$, $\cdots$, $J^{(n)}$ arranged in parallel, with a fraction $p_k$ of the light going through the $k^{\text{th}}$ Jones system $J^{(k)}$, and all the transformed beams combined (incoherently) at the output. It can also be viewed as a fluctuating system which assumes the Jones form $J^{(k)}$ with probability $p_k$. In either case, it is clear that the Mueller matrix $M$ of this nondeterministic system and its associated hermitian matrix are *corresponding convex sums*:

$$M = \sum_{k=1}^{n} p_k M(J^k), \quad H_M = \sum_{k=1}^{n} p_k \tilde{J}^{(k)} \tilde{J}^{(k)\dagger}. \tag{2.18}$$

It is useful to denote by $\{p_k, J^{(k)}\}$ the convex sum or ensemble realization represented by Eq. (2.17), or equivalently by Eq. (2.18). Obviously, such an ensemble or convex sum realization always leads to a positive semidefinite $H_M$, and positive semidefinite $H$ alone can be realized as a convex sum of projections. Thus as an immediate, and mathematically trivial, consequence of Proposition 2 we have

**Corollary**: An optical system described by $M$ is realizable as a convex sum or ensemble $\{p_k, J^{(k)}\}$ of Jones systems iff the associated $H_M \geq 0$. If $H_M \geq 0$ the number of Jones systems, $n$, needed for such a realization satisfies $n \geq r$, where $r$ is the rank of $H_M$. There is no upper limit on $n$ if $r \geq 2$.

This corollary is physically important, and has attracted considerable attention [25, 26, 27, 28, 29, 30, 31, 32].

## B. Pre-Mueller matrices and their classification

Given a $4 \times 4$ real matrix $M$, Proposition 2 gives the necessary and sufficient condition for $M$ to arise as the Mueller matrix of some Jones system $J$. That still leaves open this more general question: how to ascertain if a given matrix $M$ is a Mueller matrix? This question has an interesting history which is surprisingly recent.

In traditional polarization optics, which is formulated for plane waves and not for beams, the state space $\Omega^{(\text{pol})}$ is the collection of all Stokes vectors:

$$
\begin{aligned}
\Omega^{(\text{pol})} &= \{ S \in \boldsymbol{R}^4 \mid S_0 > 0, \ S^T G S \geq 0 \}, \\
G &= \text{diag}\,(1, -1, -1, -1) \\
S^T G S &= S_0^2 - S_1^2 - S_2^2 - S_3^2 .
\end{aligned}
\tag{2.19}
$$

Now $G$ is the 'Lorentz metric', and thus the state space $\Omega^{(\text{pol})}$ is the positive (*solid*) light cone in $\boldsymbol{R}^4$. Since a physical Mueller matrix should necessarily map states into states, our question reduces to one of effectively characterising real linear transformations in $\boldsymbol{R}^4$ which map the positive (solid) light cone *into* itself. While $SO(3,1) \cup GSO(3,1)$, where $SO(3,1)$ is the proper orthochronous Lorentz group $SO(3,1)$, is the answer in the case of *onto* maps, the *more general into case* was raised in Ref. [33] as a serious issue in polarization optics. This issue was formulated as two simple conditions that $M$ has to meet [Eqs. (2.29), (2.31) of Ref. [33]], corresponding to the demand that the intensity and degree of polarization of the output be physical *for every input pure state*. Further, the measured Mueller matrices of Howell [34] were tested for these conditions and violation was found in excess of 10%, a magnitude considerably larger than the tolerance suggested by the reported measurements. Ref. [33] thus concluded that *the Howell system fails to map the positive light cone $\Omega^{(\text{pol})}$ into itself*, and this is *possibly the first time* that a verdict of this kind was made on some published Mueller matrices.

Subsequent progress in respect of this issue was quite rapid. In a significant step forward Givens and Kostinski [35] derived, based on an impressive analysis of the spectrum of $GM^T GM$, what appeared to be a necessary and sufficient condition for $M$ to map $\Omega^{(\text{pol})}$ into itself. They analysed the Howell system based on their own condition and concluded that their results were "*in coincidence with the negative verdict on the Howell matrix delivered in* [33]". Soon after, van der Mee [29] derived a more complete set of necessary and sufficient conditions for $M$ to map $\Omega^{(\text{pol})}$ into itself; the analysis of van der Mee too was based on the spectrum of $GM^T GM$.

Decomposition of a Mueller matrix $M$ in various product forms, to gain insight into the physical effects $M$ could have on the input polarization state, has been an activity of considerable interest [36, 37, 38, 39]. The importance of obtaining the canonical or normal forms of Mueller matrices under the double-coset transformation $M \to L_\ell M L_r$, $L_\ell, L_r \in SO(3,1)$ was motivated in Ref. [20], and it was shown that the theorem of Givens and Kostinski implied that the canonical form of every nonsingular (real) matrix $M$ which maps $\Omega^{(\text{pol})}$ into itself is diagonal; i.e., $M = L_\ell M^{(1)} L_r$ where $M^{(1)} = \text{diag}\,(d_0, d_1, d_2, d_3)$, $d_0 \geq d_1 \geq d_2 \geq |d_3|$ and $L_\ell, L_r \in SO(3,1)$. It turned out that while the result of van der Mee is essentially complete [40], that of Givens and Kostinski is incomplete. This means the diagonal form $M^{(1)}$ of Ref. [20] noted above is not the only canonical form for a nonsingular $M$ mapping $\Omega^{(\text{pol})}$ into itself; there exists a non-diagonal canonical form $M^{(2)}$, and this is the case that was missed by the 'theorem' of Givens and Kostinski quoted above. In a remarkably impressive and detailed study Rao et al. [40, 41] have further explored and completed the analysis of van der Mee, leading to a complete solution to the question of canonical form for Mueller matrices, under double-coseting by $SO(3,1)$ elements, raised in Ref. [20].

Since these canonical forms play a key role in our analysis below, we list them here in a concise form. Matrices $M$

which map the state space $\Omega^{(\mathrm{pol})}$ into itself divide into two major and two minor families:

$$\text{Type I}: \quad M = L_\ell M^{(1)} L_r, \quad L_\ell,\, L_r \in SO(3,1),$$
$$M^{(1)} = \mathrm{diag}\,(d_0,\, d_1,\, d_2,\, d_3),$$
$$d_0 \geq d_1 \geq d_2 \geq |d_3|\,;$$
$$\text{Type II}: \quad M = L_\ell M^{(2)} L_r, \quad L_\ell,\, L_r \in SO(3,1),$$
$$M^{(2)} = \begin{bmatrix} d_0 & d_0 - d_1 & 0 & 0 \\ 0 & d_1 & 0 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix},$$
$$d_0 > d_1 > 0, \ \sqrt{d_0 d_1} \geq d_2 \geq |d_3|\,;$$
$$\text{Polarizer}: \quad M = L_\ell M^{(\mathrm{pol})} L_r, \quad L_\ell,\, L_r \in SO(3,1),$$
$$M^{(\mathrm{pol})} = \begin{bmatrix} d_0 & d_0 & 0 & 0 \\ d_0 & d_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ d_0 > 0\,;$$
$$\text{Pin Map}: \quad M = L_\ell M^{(\mathrm{pin})} L_r, \quad L_\ell,\, L_r \in SO(3,1),$$
$$M^{(\mathrm{pin})} = \begin{bmatrix} d_0 & 0 & 0 & 0 \\ d_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ d_0 > 0\,. \tag{2.20}$$

Since elements of $SO(3,1)$ have unit determinant, it follows that $d_3$ in the Type-I and Type-II cases is positive, negative, or zero according as $\det M$ is positive, negative, or zero. The $M$ matrices in the third and fourth families are manifestly singular. The third family is a Jones system, the associated $H$ matrix being a projection; indeed, $M^{(\mathrm{pol})}$ corresponds to a Jones matrix $J$ whose only nonvanishing element is $J_{11} = \sqrt{2d_0}$. Finally, the PinMap family is named so because $M^{(\mathrm{pin})}$ produces a *fixed* output polarization state independent of the input, the intensity of the output being independent of the state of polarization of the input. This may be contrasted with $M^{(\mathrm{pol})}$: while the output in the case of $M^{(\mathrm{pol})}$ has an input-independent state of polarization, the intensity does depend on the state of polarization of the input.

The matrix $M^{(\mathrm{pin})}$ is not a Jones system, but it is a convex sum of such systems. To see this, note that a *perfect depolarizer* represented by the Mueller matrix

$$M^{(\mathrm{depol})} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.21}$$

is a convex sum of Jones systems; it can be realized, for instance, as equal mixture of systems with Jones matrices $\tau_a$, $a = 0,\, 1,\, 2,\, 3$. That $M^{(\mathrm{pin})}$ is a convex sum of Jones systems follows from $M^{(\mathrm{pin})} = M^{(\mathrm{pol})} M^{(\mathrm{depol})}$. Alternatively, it is readily seen that $H_{M^{(\mathrm{pin})}}$ is an equal sum of two projections, and hence $M^{(\mathrm{pin})}$ is an equal mixture of the Jones systems

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.22}$$

While $M^{(\mathrm{pin})}$ is realized as convex sum of two Jones systems, $M^{(\mathrm{depol})}$ cannot be so realized with less than four Jones matrices. This follows from the fact that $H_{M^{(\mathrm{depol})}}$ is of full rank whereas $H_{M^{(\mathrm{pin})}}$ is of rank two.

The classification of canonical forms for $M$ matrices as given in (2.20) is complete in the following sense.

**Proposition 3**: Every $M$ matrix which maps the state space $\Omega^{(\mathrm{pol})}$ into itself falls *uniquely* in one of the four families described in (2.20).

That brings us to the *main thesis* of the present paper. A $4 \times 4$ real matrix $M$ will have to map the state space $\Omega^{(\text{pol})}$ into itself in order that it qualifies to be the Mueller matrix of some physical system. This is certainly a necessary condition. And, within the conventional Mueller-Stokes formalism, no conceivable further demand can be imposed on $M$. But the action of the *transversely homogeneous system represented by the numerical matrix $M$* can be extended from plane waves to paraxial beams; naturally, $M$ will then affect only the polarization degree of freedom and *act as identity on the (transverse) spatial degrees of freedom*. If $M$ indeed represented a physical system, *even this extended action should map physical states into physical states*. It turns out that this trivial looking extension is not all that trivial: there are $M$ matrices which appear physical at the level of the (restricted) state space $\Omega^{(\text{pol})}$, but fail to be physical on the extended state space. Our task in the rest of the paper is to identify precisely those $M$ matrices whose action is physical even on the extended state space. *Since only those $M$ matrices which pass this further hurdle can be called physical Mueller matrices, and pending determination of the precise demand this hurdle places on $M$, the $M$ matrices which map $\Omega^{(\text{pol})}$ into itself will be called pre-Mueller matrices.* We may thus conclude this Section by saying that Eq. (2.20) gives a complete classification of pre-Mueller matrices and their orbit structure under double-coseting by elements of $SO(3,1)$; the physical/unphysical divide of pre-Mueller matrices remains to be accomplished. This divide will be presented in Section 4 after some further preparation in Section 3.

## III.   FROM PLANE WAVES TO BEAMS: THE BCP MATRIX

We will now go beyond plane waves and consider paraxial electromagnetic beams. The simplest (quasi-) monochromatic beam field has, in a transverse plane $z = $ constant described by coordinates $(x, y) \equiv \boldsymbol{\rho}$, the form $\boldsymbol{E}(\boldsymbol{\rho}) = (E_1 \hat{\boldsymbol{x}} + E_2 \hat{\boldsymbol{y}}) \, \psi(\boldsymbol{\rho})$, where $E_1$, $E_2$ are complex constants, and the scalar-valued function $\psi(\boldsymbol{\rho})$ may be assumed to be square-integrable over the transverse plane: $\psi(\boldsymbol{\rho}) \in L^2(\boldsymbol{R}^2)$. It is clear that the polarization part $(E_1 \hat{\boldsymbol{x}} + E_2 \hat{\boldsymbol{y}})$ and the spatial dependence or modulation part $\psi(\boldsymbol{\rho})$ of such a beam are well separated, allowing one to focus attention on one aspect at a time. When one is interested in only the modulation aspect, the part $(E_1 \hat{\boldsymbol{x}} + E_2 \hat{\boldsymbol{y}})$ may be suppressed, thus leading to 'scalar optics': this is the domain of traditional Fourier optics [42]. [Fourier optics for electromagnetic beams requires a more delicate formalism [43].] On the other hand, if the spatial part $\psi(\boldsymbol{\rho})$ is suppressed we are led to the traditional polarization optics (of plane waves) described in the previous Section.

Beams whose polarization and spatial modulation separate in the above manner will be called *elementary beams*. It is clear that elementary beams remain elementary under the action of transversely homogeneous anisotropic systems like waveplates and polarizers. That they remain elementary under the action of isotropic or polarization insensitive systems like free propagation and lenses is also clear.

Now suppose we superpose or add two such elementary beam fields $(a\hat{\boldsymbol{x}} + b\hat{\boldsymbol{y}}) \, \psi(\boldsymbol{\rho})$ and $(c\hat{\boldsymbol{x}} + d\hat{\boldsymbol{y}}) \, \chi(\boldsymbol{\rho})$. The result is not of the elementary form $(e\hat{\boldsymbol{x}} + f\hat{\boldsymbol{y}}) \, \phi(\boldsymbol{\rho})$, for any $e$, $f$, $\phi(\boldsymbol{\rho})$, unless either $(a, b)$ is proportional to $(c, d)$ so that one gets committed to a *common* polarization, or $\psi(\boldsymbol{\rho})$ and $\chi(\boldsymbol{\rho})$ are proportional so that one gets committed to a *fixed* spatial mode. In other words, *the set of elementary fields is not closed under superposition*.

*Since one cannot possibly give up superposition principle in optics, one needs to go beyond the set of elementary fields and pay attention to the consequences of inseparability or entanglement of polarization and spatial variation (modulation).* We are thus led to consider (in a transverse plane) beam fields of the more general form $\boldsymbol{E}(\boldsymbol{\rho}) = E_1(\boldsymbol{\rho})\hat{\boldsymbol{x}} + E_2(\boldsymbol{\rho})\hat{\boldsymbol{y}}$. *This form is obviously closed under superposition.* We may write $\boldsymbol{E}(\boldsymbol{\rho})$ as a *generalised Jones vector*

$$\boldsymbol{E}(\boldsymbol{\rho}) \;=\; E_1(\boldsymbol{\rho})\hat{\boldsymbol{x}} + E_2(\boldsymbol{\rho})\hat{\boldsymbol{y}} \Leftrightarrow \boldsymbol{E}(\boldsymbol{\rho}) = \left[ \begin{array}{c} E_1(\boldsymbol{\rho}) \\ E_2(\boldsymbol{\rho}) \end{array} \right],$$

$$E_1(\boldsymbol{\rho}), \ E_2(\boldsymbol{\rho}) \in L^2(\boldsymbol{R}^2). \tag{3.1}$$

The intensity at location $\boldsymbol{\rho}$ corresponds to $|E_1(\boldsymbol{\rho})|^2 + |E_2(\boldsymbol{\rho})|^2$. This field is of the elementary or separable form iff $E_1(\boldsymbol{\rho})$ and $E_2(\boldsymbol{\rho})$ are linearly dependent (proportional to one another). Otherwise, polarization and spatial modulation are inseparably entangled.

The point is that the set of possible beam fields in a transverse plane constitutes the *tensor product* space $\mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$, whereas the set of all elementary fields constitutes just the *set product* $\mathcal{C}^2 \times L^2(\boldsymbol{R}^2)$ of $\mathcal{C}^2$ and $L^2(\boldsymbol{R}^2)$. [Recall that the tensor product of two vector spaces is the closure of their set product under superposition.] Thus the set product $\mathcal{C}^2 \times L^2(\boldsymbol{R}^2)$ forms a *measure zero subset* of the tensor product $\mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$. In other words, in a beam field represented by a *generic element* of $\mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$ polarization and spatial modulation should be expected to be entangled: *Entanglement is not an exception; it is the rule in* $\mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$, the space of pure states appropriate for electromagnetic beams.

If a beam described by generalised Jones vector $E(\boldsymbol{\rho})$ with $x$, $y$ components $E_1(\boldsymbol{\rho})$, $E_2(\boldsymbol{\rho})$ is passed through an $x$-polarizer, it is not only that the output will be $x$-polarized, it is certain to be in the spatial mode $E_1(\boldsymbol{\rho})$ as well. Similar conclusion holds if the beam is passed through a $y$-polarizer. Thus a (transversely homogeneous) polarizer, whose action is $\boldsymbol{\rho}$-independent, *not only chooses a polarization state, but acts as a spatial mode selector as well.* This is true even if $E_1(\boldsymbol{\rho})$ and $E_2(\boldsymbol{\rho})$ are not spatially orthogonal modes. In a similar manner, *a spatial mode selector insensitive to polarization will end up acting also as a polarization discriminater.* This is but one ramification of inseparability or entanglement between polarization and spatial modulation.

Now to handle fluctuating beams, we pass on to the beam-coherence-polarization (BCP) matrix $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}') \equiv \langle \boldsymbol{E}(\boldsymbol{\rho}) \boldsymbol{E}(\boldsymbol{\rho}')^\dagger \rangle$, defined as the ensemble average of an outer-product of (generalised) Jones vectors [5, 6]. As the name suggests, the BCP matrix describes both the coherence and polarization properties of the beam under consideration. It is a generalization of the numerical coherency matrix of plane waves considered in the previous Section, Eq. (2.2), now to the case of beam fields. It can equally well be viewed as a generalization of the mutual coherence function of scalar statistical optics to include polarization. We are free to view the BCP matrix either as the $2 \times 2$ matrix of two-point functions

$$\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}') = \begin{bmatrix} \langle E_1(\boldsymbol{\rho}) E_1(\boldsymbol{\rho}')^* \rangle & \langle E_1(\boldsymbol{\rho}) E_2(\boldsymbol{\rho}')^* \rangle \\ \langle E_2(\boldsymbol{\rho}) E_1(\boldsymbol{\rho}')^* \rangle & \langle E_2(\boldsymbol{\rho}) E_2(\boldsymbol{\rho}')^* \rangle \end{bmatrix}, \tag{3.2}$$

or as the associated column vector $\tilde{\Phi}(\boldsymbol{\rho}; \boldsymbol{\rho}')$ of two-point functions: $\tilde{\Phi}(\boldsymbol{\rho}; \boldsymbol{\rho}') = \langle \boldsymbol{E}(\boldsymbol{\rho}) \otimes \boldsymbol{E}(\boldsymbol{\rho}')^* \rangle$. For our present purpose, there is no need to make any finer distinction between the space-time and space-frequency descriptions. The two-point functions appearing in the BCP matrix may be viewed either as equal-time coherence functions or correlation functions at a particular frequency.

It is clear from the very definition (3.2) of BCP matrix that this matrix kernel, viewed as an operator from $\mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2) \to \mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$, is hermitian and positive semidefinite:

$$\Phi_{jk}(\boldsymbol{\rho}; \boldsymbol{\rho}') = \Phi_{kj}(\boldsymbol{\rho}'; \boldsymbol{\rho})^*, \quad j, k = 1, 2;$$

$$\int d^2\boldsymbol{\rho} \, d^2\boldsymbol{\rho}' \, \boldsymbol{E}(\boldsymbol{\rho})^\dagger \Phi(\boldsymbol{\rho}; \boldsymbol{\rho}') \boldsymbol{E}(\boldsymbol{\rho}') \geq 0,$$

$$\text{i.e.,} \sum_{jk} \int d^2\boldsymbol{\rho} \, d^2\boldsymbol{\rho}' E_j(\boldsymbol{\rho})^* \Phi_{jk}(\boldsymbol{\rho}; \boldsymbol{\rho}') E_k(\boldsymbol{\rho}') \geq 0,$$

$$\forall \boldsymbol{E}(\boldsymbol{\rho}) \in \mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2). \tag{3.3}$$

The positivity requirement thus demands that the expectation value of $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ be nonnegative for every Jones vector $\boldsymbol{E}(\boldsymbol{\rho})$. Hermiticity and positivity are the *defining properties* of the BCP matrix: *every $2 \times 2$ matrix of two-point functions $\Phi_{jk}(\boldsymbol{\rho}; \boldsymbol{\rho}')$ meeting just these two conditions is a valid BCP matrix of some beam of light.*

We can use the BCP matrix to define, in an obvious manner, the generalised Stokes vector $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$ [14]:

$$\begin{aligned} \Phi(\boldsymbol{\rho}; \boldsymbol{\rho}') &= \frac{1}{2} \sum_{a=0}^{3} S_a(\boldsymbol{\rho}; \boldsymbol{\rho}') \tau_a \\ &\Leftrightarrow \quad S_a(\boldsymbol{\rho}; \boldsymbol{\rho}') = \text{tr}(\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}') \tau_a). \end{aligned} \tag{3.4}$$

That this is an *invertible* relation shows that $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$ carry identical information: *action of an optical system on one defines a unique equivalent action on the other.* The hermiticity and positivity requirement on the

BCP matrix can be easily transcribed into corresponding requirements on $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$. Hermiticity reads

$$S_a(\boldsymbol{\rho}; \boldsymbol{\rho}') = S_a(\boldsymbol{\rho}'; \boldsymbol{\rho})^*, \quad a = 0, 1, 2, 3; \tag{3.5}$$

whereas positivity reads

$$\sum_{a=0}^{3} g_a \int d^2\boldsymbol{\rho} \, d^2\boldsymbol{\rho}' \, S_a(\boldsymbol{\rho}; \boldsymbol{\rho}') \widehat{S}_a(\boldsymbol{\rho}'; \boldsymbol{\rho}) \geq 0, \tag{3.6}$$

for every Stokes vector $\widehat{S}(\boldsymbol{\rho}; \boldsymbol{\rho}')$ arising from Jones vectors of the form $\boldsymbol{E}(\boldsymbol{\rho}) \in \mathcal{C}^2 \otimes L^2(\boldsymbol{R}^2)$. The signatures $g_a$ correspond to the 'Lorentz metric': $g_0 = 1$, $g_a = -1$ for $a \neq 1$.

## IV.   FROM PRE-MUELLER MATRICES TO MUELLER MATRICES: THE ROLE OF ENTANGLEMENT

We now have at our disposal all the tools necessary to determine if a given pre-Mueller matrix is a physical Mueller matrix or not. Let us consider the transformation of the generalized Stokes vector $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and the associated BCP matrix $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ under the action of a transversely homogeneous optical system described by pre-Mueller matrix $M$. We begin our analysis with pre-Mueller matrices of Type-I.

### A.   Type-I pre-Mueller matrices

We will first study pre-Mueller matrices presented in the canonical form $M^{(1)} = \mathrm{diag}\,(d_0,\, d_1,\, d_2,\, d_3)$. Extension of the conclusions to pre-Mueller matrices not in the canonical form will turn out to be quite straightforward. In view of the system's homogeneity, the action of $M^{(1)}$ is *necessarily independent* of $\boldsymbol{\rho}$, $\boldsymbol{\rho}'$, and we have

$$
M^{(1)} = \mathrm{diag}\,(d_0,\, d_1,\, d_2,\, d_3):
$$
$$
\begin{bmatrix} S_0(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_2(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_3(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}
\rightarrow
\begin{bmatrix} S_0'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_1'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_2'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_3'(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}
=
\begin{bmatrix} d_0\, S_0(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_1\, S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_2\, S_2(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_3\, S_3(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}.
$$

$$\tag{4.1}$$

The elements of the output BCP matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ associated with the output Stokes vector $S'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ resulting from the action of $M^{(1)}$ on $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$, are easily computed using Eq. (3.4):

$$
\begin{aligned}
\Phi_{11}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_0 + d_1)\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0 - d_1)\Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2, \\
\Phi_{22}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_0 + d_1)\Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0 - d_1)\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2, \\
\Phi_{12}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_2 + d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2 - d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2, \\
\Phi_{21}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_2 + d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2 - d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2.
\end{aligned}
$$

$$\tag{4.2}$$

Clearly, a *necessary condition* for the pre-Mueller matrix $M^{(1)} = \mathrm{diag}(d_0,\, d_1,\, d_2,\, d_3)$ to be a physical Mueller matrix is that the output $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ in Eq. (4.2) be a valid BCP matrix, for every valid input BCP matrix $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$. Hermiticity of $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ is manifest in view of that of $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and reality of the parameters $d_a$. Thus what remains to be checked is the positivity of $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$. While testing positivity of a generic matrix kernel could be a formidable task in general, it turns out that this test can be carried out fairly easily in the present case.

Let us take as the input a special pure state BCP matrix $\Phi^{(0)}(\boldsymbol{\rho}; \boldsymbol{\rho}') = \boldsymbol{E}(\boldsymbol{\rho})\boldsymbol{E}(\boldsymbol{\rho}')^\dagger$, corresponding to the generalised Jones vector $\boldsymbol{E}(\boldsymbol{\rho})$ which is an equal superposition of an $x$-polarized mode and a $y$-polarized mode, *the two modes*

*being spatially orthogonal*:

$$\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \boldsymbol{E}(\boldsymbol{\rho})\boldsymbol{E}(\boldsymbol{\rho}')^{\dagger},$$

$$\boldsymbol{E}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_1(\boldsymbol{\rho}) \\ \psi_2(\boldsymbol{\rho}) \end{bmatrix}, \qquad \int \psi_j(\boldsymbol{\rho})\psi_k(\boldsymbol{\rho})^* d^2\boldsymbol{\rho} = \delta_{jk}. \tag{4.3}$$

$[\psi_1(\boldsymbol{\rho})$ and $\psi_2(\boldsymbol{\rho})$ could, for instance, be two distinct Hermite-Gaussian modes.] This means that the entries of the input BCP matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ have the deterministic form $\Phi^{(0)}_{jk}(\boldsymbol{\rho};\boldsymbol{\rho}') = \psi_j(\boldsymbol{\rho})\psi_k(\boldsymbol{\rho}')^*$. A consequence of this simple (product) form is that *the four entries of the BCP matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ form an orthonormal set of (two-point) functions*:

$$\int d^2\boldsymbol{\rho}\, d^2\boldsymbol{\rho}'\, \Phi^{(0)}_{ij}(\boldsymbol{\rho};\boldsymbol{\rho}')\Phi^{(0)}_{kl}(\boldsymbol{\rho};\boldsymbol{\rho}')^* = \delta_{ik}\delta_{jl}. \tag{4.4}$$

This fact will prove to be of much value in our analysis below.

To test positivity of the output BCP matrix $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$, given in Eq. (4.2) and resulting from input $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \boldsymbol{E}(\boldsymbol{\rho})\boldsymbol{E}(\boldsymbol{\rho}')^{\dagger}$, let us define four (generalized) Jones vectors:

$$\boldsymbol{E}^{(\pm)}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_1(\boldsymbol{\rho}) \\ \pm\psi_2(\boldsymbol{\rho}) \end{bmatrix}, \quad \boldsymbol{F}^{(\pm)}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_2(\boldsymbol{\rho}) \\ \pm\psi_1(\boldsymbol{\rho}) \end{bmatrix}. \tag{4.5}$$

[The input Jones vector $\boldsymbol{E}(\boldsymbol{\rho})$ happens to coincide with $\boldsymbol{E}^{(+)}(\boldsymbol{\rho})$.] Expectation values of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ for the four Jones vectors $\boldsymbol{E}^{(\pm)}(\boldsymbol{\rho})$, $\boldsymbol{F}^{(\pm)}(\boldsymbol{\rho})$ are easily computed using Eqs. (4.2), (4.3), (4.4) and (4.5) in Eq. (3.3). These expectation values are $(d_0 + d_1) \pm (d_2 + d_3)$ for $\boldsymbol{E}^{(\pm)}(\boldsymbol{\rho})$ and $(d_0 - d_1) \pm (d_2 - d_3)$ for $\boldsymbol{F}^{(\pm)}(\boldsymbol{\rho})$.

Now positivity of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ requires, as a necessary condition, that these four expectation values be nonnegative, and this demand places on the parameters $d_a$ the constraints

$$\begin{aligned} -d_1 - d_2 - d_3 &\le d_0, \\ -d_1 + d_2 + d_3 &\le d_0; \\ d_1 + d_2 - d_3 &\le d_0, \\ d_1 - d_2 + d_3 &\le d_0. \end{aligned} \tag{4.6}$$

Violation of any one of these four conditions will render the output $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ unphysical as BCP matrix. Since the input BCP matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ is obviously physical, this will in turn render $M^{(1)}$ unphysical as Mueller matrix: Eq. (4.6) is thus a *set of necessary conditions* for the pre-Mueller matrix $M^{(1)}$ to be a Mueller matrix.

Suppose these four inequalities are met. Can we then conclude that the pre-Mueller matrix $M^{(1)}$ is a physically acceptable Mueller matrix? To answer this question in the affirmative we write in detail the associated hermitian matrix $H_{M^{(1)}} = \frac{1}{2}\sum_a d_a\,\tau_a \otimes \tau_a^* = \sum_a d_a U_{aa}$:

$$H_M = \frac{1}{2}\begin{bmatrix} d_0 + d_1 & 0 & 0 & d_2 + d_3 \\ 0 & d_0 - d_1 & d_2 - d_3 & 0 \\ 0 & d_2 - d_3 & d_0 - d_1 & 0 \\ d_2 + d_3 & 0 & 0 & d_0 + d_1 \end{bmatrix}. \tag{4.7}$$

Validity of the four inequalities in Eq. (4.6) implies, fortunately, that this matrix is positive semidefinite. This in turn implies that the given diagonal system $M^{(1)}$ is a convex sum of Jones systems, and therefore takes every BCP matrix into a BCP matrix, showing that Eq. (4.6) is *sufficient condition* for $M^{(1)}$ to be a Mueller matrix. We have thus proved

**Proposition** 4: The pre-Mueller matrix $M^{(1)} = \mathrm{diag}\,(d_0, d_1, d_2, d_3)$ is a Mueller matrix iff the associated hermitian matrix $H_{M^{(1)}} \ge 0$. That is, iff $M^{(1)}$ can be realized as a convex sum of Jones systems or, equivalently, iff the entries of $M^{(1)}$ respect the inequalities in (4.6).

Having settled the diagonal case, we now go beyond and consider non-diagonal Type-I pre-Mueller matrices. As noted in (2.20), these are necessarily of the general form $M = L_\ell M^{(1)} L_r$, where $L_\ell, L_r \in SO(3,1)$ and $M^{(1)}$ is

diagonal. We have already noted that $L_\ell$, $L_r$ are physical Mueller matrices: indeed, they correspond to deterministic systems with respective Jones matrices $J_\ell$, $J_r \in SL(2, C)$. Thus if $M^{(1)}$ is a Mueller matrix, then it has a convex sum realization $\{\, p_k, J^{(k)}\,\}$. This implies that $M = L_\ell M^{(1)} L_r$ has the convex sum realization $\{\, p_k, J_\ell J^{(k)} J_r\,\}$, and hence is a valid Mueller matrix. The converse follows by virtue of the *invertibility* of $J_\ell$, $J_r$, and we have

**Proposition** 5 : A Type-I pre-Mueller matrix, which is necessarily of the form $M = L_\ell M^{(1)} L_r$ with $L_\ell$, $L_r \in SO(3, 1)$ and $M^{(1)}$ diagonal, is a physical Mueller matrix iff $M^{(1)}$ is.

## B. Type-II pre-Mueller matrices

Having fully classified Type-I pre-Mueller matrices into Mueller and non-Mueller matrices, we now turn our attention to Type-II pre-Mueller matrices. The analysis turns out to be quite parallel to the one in the previous sub-Section.

Recall from Section 2 that a Type-II pre-Mueller matrix in its canonical form $M^{(2)}$ has only one nonvanishing off-diagonal element whose value is fixed by the diagonals, namely $M_{01} = d_0 - d_1$, where $d_0$, $d_1$, $d_2$, $d_3$ are the diagonals. The action of $M^{(2)}$ on $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ can be computed as before. The (generalised) Stokes vector has this simple transformation law:

$$
M^{(2)} : \quad S(\boldsymbol{\rho}; \boldsymbol{\rho}') \to S'(\boldsymbol{\rho}; \boldsymbol{\rho}') = \begin{bmatrix} S_0'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_1'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_2'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_3'(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}
$$
$$
= \begin{bmatrix} d_0 S_0(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0 - d_1) S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_1 S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_2 S_2(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_3 S_3(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}. \tag{4.8}
$$

The elements of the output BCP matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ associated with $S'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and computed from (3.4) are

$$
\begin{aligned}
\Phi_{11}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= d_0\, \Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,, \\
\Phi_{22}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= d_1\, \Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0 - d_1)\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,, \\
\Phi_{12}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_2 + d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2 - d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2, \\
\Phi_{21}'(\boldsymbol{\rho}; \boldsymbol{\rho}') &= [\,(d_2 + d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2 - d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}')\,]/2.
\end{aligned} \tag{4.9}
$$

As in the case of $M^{(1)}$, the canonical form pre-Mueller matrix $M^{(2)}$ does not couple the pair $\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}')$, $\Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}')$ with $\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}')$, $\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}')$.

Again, a necessary condition for the pre-Mueller matrix $M^{(2)}$ to be a physically acceptable Mueller matrix is that the output $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ in Eq. (4.9) be a valid BCP matrix for every valid input BCP matrix $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$. As in the case of $M^{(1)}$, let us take as input the pure state BCP matrix $\Phi^{(0)}(\boldsymbol{\rho}; \boldsymbol{\rho}') = \boldsymbol{E}(\boldsymbol{\rho})\boldsymbol{E}(\boldsymbol{\rho}')^\dagger$, with $\boldsymbol{E}(\boldsymbol{\rho})$ as described in Eq. (4.3). To test positivity of the output BCP matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$, given in Eq. (4.9) and resulting from input $\Phi^{(0)}(\boldsymbol{\rho}; \boldsymbol{\rho}') = \boldsymbol{E}(\boldsymbol{\rho})\boldsymbol{E}(\boldsymbol{\rho}')^\dagger$, we use in place of $\boldsymbol{E}^{(\pm)}(\boldsymbol{\rho})$ and $\boldsymbol{F}^{(\pm)}(\boldsymbol{\rho})$ slightly modified (generalized) Jones vectors $\boldsymbol{E}^{(\theta)}(\boldsymbol{\rho})$ and $\boldsymbol{F}^{(\theta)}(\boldsymbol{\rho})$ :

$$
\boldsymbol{E}^{(\theta)}(\boldsymbol{\rho}) = \begin{bmatrix} \cos\theta\, \psi_1(\boldsymbol{\rho}) \\ \sin\theta\, \psi_2(\boldsymbol{\rho}) \end{bmatrix}, \quad \boldsymbol{F}^{(\theta)}(\boldsymbol{\rho}) = \begin{bmatrix} \cos\theta\, \psi_2(\boldsymbol{\rho}) \\ \sin\theta\, \psi_1(\boldsymbol{\rho}) \end{bmatrix}. \tag{4.10}
$$

Expectation values of the output BCP matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ for these two families of Jones vectors can be computed as before. These expectation values are $d_0 \cos^2\theta + d_1 \sin^2\theta + (d_2 + d_3)\cos\theta\sin\theta$ for $\boldsymbol{E}^\theta(\boldsymbol{\rho})$ and $(d_0 - d_1)\sin^2\theta + (d_2 - d_3)\cos\theta\sin\theta$ for $\boldsymbol{F}^\theta(\boldsymbol{\rho})$.

Now positivity of $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ requires, as a necessary condition, that these expectation values be nonnegative for all $0 \le \theta < \pi$, and this requirement is seen to be equivalent to the pair of conditions $d_0 d_1 \ge (d_2 + d_3)^2/4$, $d_2 - d_3 = 0$;

these arise respectively from the $\boldsymbol{E}^{(\theta)}$ and $\boldsymbol{F}^{(\theta)}$ families. We may rewrite these as

$$d_3 = d_2, \quad (d_2)^2 \le d_0 d_1. \tag{4.11}$$

This is a pair of *necessary conditions* for the pre-Mueller matrix $M^{(2)}$ to be a Mueller matrix. The condition $(d_2)^2 \le d_0 d_1$ is already part of the definition of $M^{(2)}$, and thus $d_3 = d_2$ is the new requirement arising from consideration of the action of $M^{(2)}$ on BCP matrices, i.e., from consideration of entanglement.

Our next task is to show that these conditions are sufficient as well. To this end we proceed as in the case of $M^{(1)}$ and compute the hermitian matrix $H_{M^{(2)}}$ associated with $M^{(2)}$:

$$H_{M^{(2)}} = \begin{bmatrix} d_0 & 0 & 0 & \frac{1}{2}(d_2 + d_3) \\ 0 & 0 & \frac{1}{2}(d_2 - d_3) & 0 \\ 0 & \frac{1}{2}(d_2 - d_3) & d_0 - d_1 & 0 \\ \frac{1}{2}(d_2 + d_3) & 0 & 0 & d_1 \end{bmatrix}. \tag{4.12}$$

The inequalities in Eq. (4.11) are precisely the conditions under which $H_{M^{(2)}}$ is positive. This in turn implies that $M^{(2)}$ satisfying (4.11) is a convex sum of Jones systems, and hence is a Mueller matrix. We have thus proved

**Proposition 6**: The pre-Mueller matrix $M^{(2)}$ is a Mueller matrix iff the associated hermitian matrix $H_{M^{(2)}} \ge 0$. That is, iff $M^{(2)}$ can be realized as a convex sum of Jones systems or, equivalently, iff the entries of $M^{(2)}$ respect the inequalities in (4.11).

We can now proceed to consider Type-II pre-Mueller matrices which are not of the canonical form $M^{(2)}$. We know from Section 2 that any such matrix has the form $M = L_\ell M^{(2)} L_r$, where $L_\ell, L_r \in SO(3,1)$. By considerations similar to the ones leading to Proposition 6 in the Type-I case, we arrive at

**Proposition 7**: A type-II pre-Mueller matrix, which is necessarily of the form $M = L_\ell M^{(2)} L_r$ with $L_\ell, L_r \in SO(3,1)$, is a Mueller matrix iff $M^{(2)}$ is.

Having completed classification of the pre-Mueller matrices in the Type-I and Type-II families into physical and non-physical ones, we are now left with two minor families to handle. As noted following (2.20), $M^{(\text{pol})}$ is a Jones system. Let $J$ be the Jones matrix representing this system (polariser). In view of the two-to-one homomorphism between $SL(2, C)$ and $SO(3,1)$ alluded to earlier, $L_\ell, L_r \in SO(3,1)$ define respective Jones matrices $J_\ell, J_r$ of unit determinant; these Jones matrices are *unique except for multiplicative factor* $\pm 1$ and, as is well known, this signature ambiguity is of nontrivial origin. Thus $M = L_\ell M^{(\text{pol})} L_r$ is a Jones system with Jones matrix $\pm J_\ell J J_r$, and hence is physical. Similar argument will show that the last family, namely the PinMap family, too has no non-physical $M$ matrix. For completeness, we state the situation in respect of these two minor families as the following

**Proposition 8**: Pre-Mueller matrices belonging to the Polarizer and Pin Map families are respectively Jones systems and convex sums of Jones systems. Their associated $H$ matrices are positive semidefinite, and all pre-Mueller matrices in these two families are physical Mueller matrices.

## C. Complete Characterization of Mueller Matrices

In the last two sub-Sections we carried out a complete classification of pre-Mueller matrices into physical and non-physical ones. Double-coseting under the $SO(3,1)$ group has played such an important role in this process that we capture this role as a separate result.

**Proposition 9**: Given two $4 \times 4$ real matrices $M$ and $M'$ which are in the same double-coset orbit under $SO(3,1)$, i.e., $M' = L_\ell M L_r$ for some $L_\ell, L_r \in SO(3,1)$, $M'$ is a convex sum of Jones systems iff $M$ is. And $H_{M'} \ge 0$ iff $H_M \ge 0$. In other words, $M'$ is a Mueller matrix iff $M$ is.

*Proof*: Suppose $M$ has the convex sum realization $\{p_k, J^{(k)}\}$, i.e., $M = \sum_k p_k M(J^{(k)})$. Then, clearly, $M'$ has the convex sum realization $\{p_k, J_\ell J^{(k)} J_r\}$. Conversely, if $M'$ has the convex sum realization $\{p'_k, J'^{(k)}\}$, then $M$ has the convex sum realization $\{p'_k, (J_\ell)^{-1} J'^{(k)} (J_r)^{-1}\}$.

Now suppose $H_M \geq 0$. This means $H_M = \sum_k p_k \, \tilde{J}^{(k)} \, \tilde{J}^{(k)\,\dagger}$, $p_k > 0$. This immediately implies $H_{M'} = \sum_k p_k \, (J_\ell \widetilde{J^{(k)}} J_r)(J_\ell \widetilde{J^{(k)}} J_r)^{\dagger}$, which proves its positivity. Here $(J_\ell \widetilde{J^{(k)}} J_r)$, as usual, denotes the column vector associated with the $2 \times 2$ matrix $J_\ell J^{(k)} J_r$. The converse follows from the invertibility of $J_\ell$, $J_r$, completing proof of the Proposition.

With this proof of the principal conclusion of this paper is complete. Our main theorem may thus be stated as follows.

**Theorem**: A $4 \times 4$ real matrix $M$ is a Mueller matrix iff the associated hermitian matrix $H_M \geq 0$. Every physically acceptable Mueller matrix is a convex sum of Mueller-Jones matrices.

### D. The role of entanglement: An illustrative example

We present a simple example to illustrate the kind or restrictions on $M$ matrices brought in by consideration of entanglement. Let us restrict attention to $M$ matrices of the special simple three-parameter form

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & d_1 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix}. \tag{4.13}$$

We are obviously in the Type-I situation, but we are not considering here the $SO(3,1)$ orbit under double-coseting.

It is clear that $M$ will map Stokes vectors into Stokes vectors if and only if $M$ satisfies the following three conditions:

$$M: \ \Omega^{(\mathrm{pol})} \ \rightarrow \ \Omega^{(\mathrm{pol})} \Leftrightarrow$$
$$-1 \leq d_k \leq 1, \quad k = 1, 2, 3. \tag{4.14}$$

*In the absence of considerations of entanglement these would have been the only conditions $M$ will need to satisfy.* Thus the allowed region in the Euclidean space $\boldsymbol{R}^3$ spanned by the parameters $(d_1, d_2, d_3)$ would have been the cube with vertices at $(\pm 1, \pm 1, \pm 1)$.

Now each one of the four conditions in (4.6) with $d_0 = 1$, arising out of consideration of entanglement, forbids the (open) half-space on one side of a plane. *Thus the allowed region is the intersection of the four allowed half-spaces.* This region is clearly the tetrahedron with vertices at $(1, 1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$ and $(-1, -1, 1)$.

Thus, it is not the entire cubical region (4.14) permitted by conventional wisdom, but only the solid tetrahedron, with one third volume of the cube, that stands the closer scrutiny presented by consideration of entanglement. This is illustrated in Fig.1. The region outside the tetrahedron is unphysical and does not correspond to Mueller matrices: points in the cube but exterior to the tetrahedron correspond to pre-Mueller matrices which are not physical Mueller matrices; those outside the cube correspond to $M$ matrices which are not even pre-Mueller matrices.

## V. CONCLUDING REMARKS

We conclude with some further observations. In the mathematics literature and in the literature of quantum information theory, what we have called pre-Mueller matrices go by the name *positive maps*, and the subset of pre-Mueller matrices which are physically acceptable in the sense of our main theorem corresponds to what are called *completely positive maps*. But we have endevoured here to arrive at a physical characterization of Mueller matrices entirely within the framework of BCP matrices familiar to the classical optics community, without resorting to the mathematical theory of these maps.

Secondly, a pre-Mueller matrix which fails our main theorem *will not* produce any unphysical effect acting on the coherency matrix of plane waves or on the BCP matrix of elementary (or polarization-modulation separable) beams. It follows that BCP matrices which are convex sums of elementary beams will not be able to witness the failure of a pre-Mueller matrix $M$ whose associated $H_M$ is not positive semidefinite. Only BCP matrices corresponding to

entangled generalized Jones vectors can expose the unphysical nature of a pre-Mueller matrix which violates our main theorem. In other words, *pre-Mueller matrices cannot be further divided into physical and unphysical subsets without consideration of entanglement.*

Finally, ever since it was proved that every Jones system corresponds to a Mueller matrix whose associated $H$ matrix is a projection [24], it has been clear that ensembles of Jones systems necessarily correspond to positive semidefinite $H$ matrices, and conversely. It has thus been occasionally suggested by various authors, beginning with [25], that considerations of Mueller matrices might be restricted to only such ensembles. But it has remained only a suggestion, and one without any physical basis, and hence could not set aside as unphysical an experimentally measured Mueller matrix whose associated $H$ matrix has a negative eigenvalue, *particularly when the reported Mueller system was not realized by the experimenter specifically as an ensemble of Jones systems.* For instance the symmetric Mueller matrix of van Zyl [44] was analysed in Ref. [20], and was found to be a matrix of Type-I, with canonical-form parameters $d_0 = 0.9735$, $d_1 = 0.9112$, $d_2 = 0.4640$, $d_3 = -0.3838$. This clearly violates [the second constraint in] Eq. (4.6) by a substantial extent. Equivalently, the eigenvalues of $H_M$ are 1.0906, 0.8393, 0.4526, $-0.3825$.

That $H_M$ in this case is not positive, and hence the van Zyl system is not a convex sum of Jones systems was always known. However, one did not have hitherto a physical basis on which this $M$ could be judged as unphysical. Non-quantum entanglement has now given us such a physical basis.

[1] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995), Chap. 6.

[2] C. Brosseau, *Fundamentals of Polarized Light: A Statistical Approach* (Wiley, N.Y., 1998).

[3] J.J. Gil, "Polarimetric characterization of light and media," Eur. Phys. J. Appl. Phys. **40**, 1–47 2007).

[4] D.F.V. James, "Change of polarization of light-beams on propagation through free space," J. Opt. Soc. Am. A **11**, 1641–1643 (1994).

[5] F. Gori, "Matrix treatment of partially polarized, partially coherent beams," Opt. Lett. **23**, 241–243 (1998).

[6] F. Gori, M. Santarsiero, S. Vicalvi, R. Borghi, and G. Gauttari, "Beam coherence-polarization matrix," J. Opt. A: Pure Appl. Opt. **7**, 941–951 (1998).

[7] F. Gori, M. Santarsiero, G. Piquero, R. Borghi, A. Mondello, and R. Simon, "Partially polarized Gaussian Schell-model beams," J. Opt. A: Pure Appl. Opt. **3**, 1–9 (2001).

[8] F. Gori, M. Santarsiero, R. Simon, G. Piquero, R. Borghi, and R. Gauttari, "Coherent-mode decomposition of partially coherent partially polarized beams," J. Opt. Soc. Am. A **20**, 78–84 (2003).

[9] E. Wolf, "Unified theory of coherence and polarization of random electromagnetic beams," Phys. Lett. A **312**, 263–267 (2003).

[10] J. Tervo, T. Setala, and A.T. Friberg, "Degree of coherence of electromagnetic fields," Opt. Express **11**, 1137–1142 (2003).

[11] T. Setala J. Tervo, and A.T. Friberg, "Complete electromagnetic coherence in the space frequency domain," Opt. Lett. **29**, 328–330 (2004).

[12] M.R. Dennis, "Geometrical interpretation of the three-dimensional coherence matrix for nonparaxial polarization," J. Opt. A: Pure Appl. Opt. **6**, S26–S31 (2004).

[13] P. Refregier and F. Goudail, "Invariant degrees of coherence of partially polarized light," Opt. Express **13**, 1051–1060 (2005).

[14] O. Korotkova and E. Wolf, "Generalized Stokes parameters of random electromagnetic beams," Opt. Lett. **30**, 198–200 (2005).

[15] F. Gori, M. Santarsiero, and R. Borghi, "Maximizing Young's fringe visibility through reversible optical transformations," Opt. Lett. **32**, 688–590 (2007).

[16] A. Luis, "Degree of coherence for vectorial electromagnetic fields as a distance between correlation matrices," J. Opt. Soc. Am. A **24**, 1063–1068 (2007).

[17] R. Matrinez-Herrero and P.M. Mejias, "Maximum visibility under unitary transformations in two-pinhole interference for electromagnetic fields," Opt. Lett. **32**, 1471–1473 (2007).

[18] R. Matrinez-Herrero and P.M. Mejias, "Relation between degrees of coherence for electromagnetic fields," Opt. Lett. **32**, 1504–1506 (2007).

[19] R. Barakat, "Bilinear constraints between elements of the $4 \times 4$ Mueller-Jones transfer matrix of polarization theory," Opt. Commun. **38**, 159–161 (1981).

[20] R. Sridhar and R. Simon, "Normal form for Mueller matrices in polarization optics," J. Mod. Opt. **41**, 1903–1915 (1994).

[21] G.G. Stokes, "On the composition and resolution of streams of polarized light from different sources," Trans. Cambridge Philos. Soc. **9**, 399–416 (1852).

[22] K.D. Abhyankar and A.L. Fymat, "Relationships between the elements of the phase matrix for scattering," J. Math. Phys. **10**, 1935–1938 (1969).

[23] E.S. Fry and E.W. Kattawar, "Relationships between elements of the Stokes matrix," Appl. Opt. **20**, 2811–2814 (1981).

[24] R. Simon, "The connection between Mueller and Jones matrices of polarization optics," Opt. Commun. **42**, 293–297 (1982).

[25] K. Kim, L. Mandel, and E. Wolf, "Relationship between Jones and Mueller matrices for random media," J. Opt. Soc. Am. A **4**, 433–437 (1987).

[26] J.J. Gil and E. Bernabeu, "A depolarization criterion in Mueller matrices," Opt. Acta **32**, 259–261 (1985).

[27] J.J. Gil and E. Bernabeu, "Obtainment of the polarizing and retardation parameters of a nondepolarizing optical system from the decomposition of its Mueller matrix," Optik **76**, 67–71 (1987).

[28] S.R. Cloude, "Conditions for the physical realizability of matrix operations in polarimetry," in *Polarization Considerations in Optical Systems*, Ed. R.A. Chipman (SPIE, Vol. 1166, 1989).

[29] V.M. van der Mee, "An eigenvalue criterion for matrices transforming Stokes parameters," J. Math. Phys. **34**, 5072–5088 (1993).

[30] D.G.M. Anderson and R. Barakat, "Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix," J. Opt. Soc. Am. A **11**, 2305–2319 (1994).

[31] A. Aiello, G. Puentes, and J.P. Woerdman, "Linear optics and quantum maps," Phys. Rev. A **76**, art. 032323 (2007).

[32] Sudha, A.V.G. Rao, A.R.U. Devi, and A.K. Rajagopal, "Positive-operator-valued measure view of the ensemble approach to polarization optics," J. Opt. Soc. Am. A **25**, 874–880 (2008).

[33] M. Sanjay Kumar and R. Simon, "Characterization of Mueller matrices in polarization optics," Opt. Commun. **88**, 464–470 (1992).

[34] B.J. Howell, "Measurement of the polarization effects of an instrument using partially polarized light," Appl. Opt. **18**, 809–812 (1979).

[35] C.R. Givens and B. Kostinski, "A simple necessary and sufficient condition on physically realizable Mueller matrices," J. Mod. Opt. **40**, 471–481 (1993).

[36] S.Y. Lu and R.A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," J. Opt. Soc. Am. A **13**, 1106–1113 (1996).

[37] R. Ossikovski, A. De Martino, and S. Guyot, "Forward and reverse product decompositions of depolarizing Mueller matrices," Opt. Lett. **32**, 689–691 (2007).

[38] V. Devlaminck and P. Terrier, "Definition of a parameter form for non-singular Mueller matrices," J. Opt. Soc. Am. A **25**, 2636–2643 (2008).

[39] F. Boulvert, G. Le Brun, B. Le Jeune, J. Cariou, and L. Martin, "Decomposition algorithm for an experimental Mueller matrix," Opt. Commun. **282**, 692–704 (2009).

[40] A.V. Gopala Rao, K.S. Mallesh, and Sudha, "On the algebraic characterization of a Mueller matrix in polarization optics: I. Identifying a Mueller matrix from its $N$ matrix," J. Mod. Opt. **45**, 955–987 (1998).

[41] A.V. Gopala Rao, K.S. Mallesh, and Sudha, "On the algebraic characterization of a Mueller matrix in polarization optics: II. Necessary and sufficient condition for Jones-derived Mueller matrix," J. Mod. Opt. **45**, 989–999 (1998).

[42] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York (1968).

[43] N. Mukunda, R. Simon, and E.C.G. Sudarshan, "Fourier optics for the Maxwell field: Formalism and applications," J. Opt. Soc. Am. A **2**, 416–426 (1985).

[44] J.J. van Zyl, C.H. Papas, and C. Elachi, "On the optimum polarization of incoherently reflected waves," IEEE Trans. Antennas Prop. **AP-35**, 818–825 (1987).